

## **STIC Biotechnology Systems Branch**

### **RAW SEQUENCE LISTING** **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/522,341  
Source: PCT  
Date Processed by STIC: 2-17-05

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 4.2.2 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):  
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/24/05

**BEST AVAILABLE COPY**



PCT

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,341

DATE: 02/17/2005

TIME: 12:04:43

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\02172005\J522341.raw

2 <110> APPLICANT: Kock, Michael  
 3 Frank, Markus  
 4 Badur, Ralf  
 6 <120> TITLE OF INVENTION: Novel Selection Processes  
 8 <130> FILE REFERENCE: 532622010400  
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/522,341  
 C--> 11 <141> CURRENT FILING DATE: 2005-01-25  
 13 <160> NUMBER OF SEQ ID NOS: 71  
 15 <170> SOFTWARE: PatentIn Ver. 2.1

## ERRORED SEQUENCES

Does Not Comply  
 Corrected Diskette Needed

(pg. 4, 6) ←

350 <210> SEQ ID NO: 4  
 351 <211> LENGTH: 427  
 352 <212> TYPE: PRT  
 353 <213> ORGANISM: Artificial sequence  
 W--> 354 <220> FEATURE: pls insert  
 354 <223> OTHER INFORMATION: Description of the artificial sequence: coding for  
 355 cytosine deaminase (codA)  
 O/E 357 <400> SEQUENCE: 4  
 358 Met Ser Asn Asn Ala Leu Gln Thr Ile Ile Asn Ala Arg Leu Pro Gly  
 359 1 5 10 15  
 361 Glu Glu Gly Leu Trp Gln Ile His Leu Gln Asp Gly Lys Ile Ser Ala  
 362 20 25 30  
 364 Ile Asp Ala Gln Ser Gly Val Met Pro Ile Thr Glu Asn Ser Leu Asp  
 365 35 40 45  
 367 Ala Glu Gln Gly Leu Val Ile Pro Pro Phe Val Glu Pro His Ile His  
 368 50 55 60  
 370 Leu Asp Thr Thr Gln Thr Ala Gly Gln Pro Asn Trp Asn Gln Ser Gly  
 371 65 70 75 80  
 373 Thr Leu Phe Glu Gly Ile Glu Arg Trp Ala Glu Arg Lys Ala Leu Leu  
 374 85 90 95  
 376 Thr His Asp Asp Val Lys Gln Arg Ala Trp Gln Thr Leu Lys Trp Gln  
 377 100 105 110  
 379 Ile Ala Asn Gly Ile Gln His Val Arg Thr His Val Asp Val Ser Asp  
 380 115 120 125  
 382 Ala Thr Leu Thr Ala Leu Lys Ala Met Leu Glu Val Lys Gln Glu Val  
 383 130 135 140  
 385 Ala Pro Trp Ile Asp Leu Gln Ile Val Ala Phe Pro Gln Glu Gly Ile  
 386 145 150 155 160  
 388 Leu Ser Tyr Pro Asn Gly Glu Ala Leu Leu Glu Glu Ala Leu Arg Leu  
 389 165 170 175

←  
 FYI:  
pls insert  
 <220> whenever  
 <221>, <222>  
 ok <223>  
 is present.

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,341

DATE: 02/17/2005

TIME: 12:04:43

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\02172005\J522341.raw

```

391 Gly Ala Asp Val Val Gly Ala Ile Pro His Phe Glu Phe Thr Arg Glu
392                180                185                190
394 Tyr Gly Val Glu Ser Leu His Lys Thr Phe Ala Leu Ala Gln Lys Tyr
395                195                200                205
397 Asp Arg Leu Ile Asp Val His Cys Asp Glu Ile Asp Asp Glu Gln Ser
398                210                215                220
400 Arg Phe Val Glu Thr Val Ala Ala Leu Ala His His Glu Gly Met Gly
401 225                230                235                240
403 Ala Arg Val Thr Ala Ser His Thr Thr Ala Met His Ser Tyr Asn Gly
404                245                250                255
406 Ala Tyr Thr Ser Arg Leu Phe Arg Leu Leu Lys Met Ser Gly Ile Asn
407                260                265                270
409 Phe Val Ala Asn Pro Leu Val Asn Ile His Leu Gln Gly Arg Phe Asp
410                275                280                285
412 Thr Tyr Pro Lys Arg Arg Gly Ile Thr Arg Val Lys Glu Met Leu Glu
413                290                295                300
415 Ser Gly Ile Asn Val Cys Phe Gly His Asp Asp Val Phe Asp Pro Trp
416 305                310                315                320
418 Tyr Pro Leu Gly Thr Ala Asn Met Leu Gln Val Leu His Met Gly Leu
419                325                330                335
421 His Val Cys Gln Leu Met Gly Tyr Gly Gln Ile Asn Asp Gly Leu Asn
422                340                345                350
424 Leu Ile Thr His His Ser Ala Arg Thr Leu Asn Leu Gln Asp Tyr Gly
425                355                360                365
427 Ile Ala Ala Gly Asn Ser Ala Asn Leu Ile Ile Leu Pro Ala Glu Asn
428                370                375                380
430 Gly Phe Asp Ala Leu Arg Arg Gln Val Pro Val Arg Tyr Ser Val Arg
431 385                390                395                400
433 Gly Gly Lys Val Ile Ala Ser Thr Gln Pro Ala Gln Thr Thr Val Tyr
434                405                410                415
436 Leu Glu Gln Pro Glu Ala Ile Asp Tyr Lys Arg
437                420                425

```

4798 &lt;210&gt; SEQ ID NO: 50

4799 &lt;211&gt; LENGTH: 27

4800 &lt;212&gt; TYPE: DNA

4801 &lt;213&gt; ORGANISM: Artificial sequence

W--&gt; 4802 &lt;220&gt; FEATURE:

4803 &lt;223&gt; OTHER INFORMATION: Description of the artificial sequence:

4804 oligonucleotide primer

4806 &lt;400&gt; SEQUENCE: 50

4807 cgtgaatacgcgcgtggagtcg

4810 &lt;210&gt; SEQ ID NO: 51

4811 &lt;211&gt; LENGTH: 26

4812 &lt;212&gt; TYPE: DNA

4813 &lt;213&gt; ORGANISM: Artificial sequence

W--&gt; 4814 &lt;220&gt; FEATURE:

4815 &lt;223&gt; OTHER INFORMATION: Description of the artificial sequence:

4816 oligonucleotide primer

4818 &lt;400&gt; SEQUENCE: 51

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,341

DATE: 02/17/2005

TIME: 12:04:44

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\02172005\J522341.raw

E--> 4819 eggcaggata atcaggttgg  
 4842 <210> SEQ ID NO: 53  
 4843 <211> LENGTH: 27  
 4844 <212> TYPE: DNA  
 4845 <213> ORGANISM: Artificial sequence

W--> 4846 <220> FEATURE:  
 4847 <223> OTHER INFORMATION: Description of the artificial sequence:  
 4848 oligonucleotide primer  
 4850 <400> SEQUENCE: 53

E--> 4851 gtcaacgtaa ccaacocctgc  
 5506 <210> SEQ ID NO: 59  
 5507 <211> LENGTH: 1011  
 5508 <212> TYPE: DNA  
 5509 <213> ORGANISM: Zea mays

W--> 5510 <220> FEATURE:  
 5511 <221> NAME/KEY: CDS  
 5512 <222> LOCATION: (1)..(981)  
 5513 <223> OTHER INFORMATION: coding for 5-methylthioribose kinase  
 5515 <400> SEQUENCE: 59

5516	gca cga gca ctc ctc tcc tct cct ctc gcc ggc gca tcg ccc gac tgt	48
5517	Ala Arg Ala Leu Leu Ser Ser Pro Leu Ala Gly Ala Ser Pro Asp Cys	
5518	1 5 10 15	
5520	cag tca gcc tca gcc atg gcc gcg gag gag gag cag ggc ttc cgc ccg	96
5521	Gln Ser Ala Ser Ala Met Ala Ala Glu Glu Glu Gln Gly Phe Arg Pro	
5522	20 25 30	
5524	ctg gac gag tcg tcc ctg ctc gcc tac atc aag gcc acg ccg gcg ctc	144
5525	Leu Asp Glu Ser Ser Leu Leu Ala Tyr Ile Lys Ala Thr Pro Ala Leu	
5526	35 40 45	
5528	gcc tcc cgc ctc ggc ggc ggt ggc agt cta gac tcc atc gag atc aag	192
5529	Ala Ser Arg Leu Gly Gly Gly Ser Leu Asp Ser Ile Glu Ile Lys	
5530	50 55 60	
5532	gag gtc ggc gac ggc aac ctc aac ttc gtc tac atc gtg cag tcc gag	240
5533	Glu Val Gly Asp Gly Asn Leu Asn Phe Val Tyr Ile Val Gln Ser Glu	
5534	65 70 75 80	
5536	gcc ggc gcc atc gtc gtc aag cag gcg ctc ccg tac gtg cgc tgc gtg	288
5537	Ala Gly Ala Ile Val Val Lys Gln Ala Leu Pro Tyr Val Arg Cys Val	
5538	85 90 95	
5540	ggg gat tcg tgg ccc atg acg cgg gag cgc gcc tac ttc gag gcc tcc	336
5541	Gly Asp Ser Trp Pro Met Thr Arg Glu Arg Ala Tyr Phe Glu Ala Ser	
5542	100 105 110	
5544	acg ctg cgg gag cac ggc cgc ctg tgc ccg gag cac acc ccc gag gtg	384
5545	Thr Leu Arg Glu His Gly Arg Leu Cys Pro Glu His Thr Pro Glu Val	
5546	115 120 125	
5548	tac cac ttc gac cgg acc ttg tcg ctg atg ggg atg cgc tac atc gag	432
5549	Tyr His Phe Asp Arg Thr Leu Ser Leu Met Gly Met Arg Tyr Ile Glu	
5550	130 135 140	
5552	ccc ccg cac atc atc ctc cgc aag ggc ctc gtc gcc ggt gtc gag tac	480
5553	Pro Pro His Ile Ile Leu Arg Lys Gly Leu Val Ala Gly Val Glu Tyr	
5554	145 150 155 160	

20

20

See  
 pg. 2  
 For error  
 explanation.

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,341

DATE: 02/17/2005

TIME: 12:04:44

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\02172005\J522341.raw

```

5556 ccg ctg ctc gcc gac cac atg tcc gat tac atg gcc aag acg ctc ttc      528
5557 Pro Leu Leu Ala Asp His Met Ser Asp Tyr Met Ala Lys Thr Leu Phe
5558              165              170              175
5560 ttc acc tcc ctc ctc tat aac aat acc acg gat cat aag aac gga gtt      576
5561 Phe Thr Ser Leu Leu Tyr Asn Asn Thr Thr Asp His Lys Asn Gly Val
5562              180              185              190
5564 gct aag tac tct gcg aac gtg gag atg tgt agg ctc acg gag caa gtt      624
5565 Ala Lys Tyr Ser Ala Asn Val Glu Met Cys Arg Leu Thr Glu Gln Val
5566              195              200              205
5568 gtg ttc tcg gac cca tac cgt gtt tcc aaa ttt aat cgg tgg acc tcg      672
5569 Val Phe Ser Asp Pro Tyr Arg Val Ser Lys Phe Asn Arg Trp Thr Ser
5570              210              215              220
5572 cct tat ctc gac aaa gat gct gag gca gtt cgc gag gat gat gag ctc      720
5573 Pro Tyr Leu Asp Lys Asp Ala Glu Ala Val Arg Glu Asp Asp Glu Leu
5574 225              230              235              240
5576 aag ttg gaa gta gct ggg ctg aaa tcg atg ttt atc gag aga gct caa      768
5577 Lys Leu Glu Val Ala Gly Leu Lys Ser Met Phe Ile Glu Arg Ala Gln
5578              245              250              255
5580 gct ctg att cat gga gat ctc cac act ggt tct atc atg gtg acc gaa      816
5581 Ala Leu Ile His Gly Asp Leu His Thr Gly Ser Ile Met Val Thr Glu
5582              260              265              270
5584 gtt caa ctc aag tca ttg atc cag aat ttg ggt tct atg ggg cca atg      864
5585 Val Gln Leu Lys Ser Leu Ile Gln Asn Leu Gly Ser Met Gly Pro Met
5586              275              280              285
5588 ggg ttt gat att ggg agc ctt cct tgg aaa cct gat ttt ggg cat act      912
5589 Gly Phe Asp Ile Gly Ser Leu Pro Trp Lys Pro Asp Phe Gly His Thr
5590              290              295              300
5592 atg cac aga atg ggc atg ctg atc aag cga atg atc gta agg ctt aca      960
5593 Met His Arg Met Gly Met Leu Ile Lys Arg Met Ile Val Arg Leu Thr
5594 305              310              315              320
5596 aga atg gat ctt gaa gac aat tgaagagtcg tggaatttgt tccacaaaaa      1011
E--> 5597 arg met asp Leu glu asp asn

```

325

← move to (325) delete

10/522,341

Page 5

&lt;210&gt; 55

&lt;211&gt; 5674

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

<223> Description of the artificial sequence: vector  
construct pBluKS-nitP-STLS1-35S-T

&lt;400&gt; 55

ccagcttttg	ttccctttag	tgagggttaa	tttcgagctt	ggcgtaatca	tggtcatagc	60
tgtttctgt	gtgaaattgt	tatccgctca	caattccaca	caacatacga	gccggaagca	120
taaagtgtaa	agcctggggg	gcctaattgag	tgagctaact	cacattaatt	gcgttgcgct	180
cactgcccgc	tttccagtcg	ggaaacctgt	cgtgccagct	gcattaatga	atcgccaac	240
gcgcggggag	aggcggtttg	cgtattgggc	gctcttccgc	ttcctcgctc	actgactcgc	300
tgcgctcgtt	cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	gtaatacggg	360
tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaggc	cagcaaaagg	420
ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	480
agcatcacia	aatcgacgc	tcaagtcaga	ggtggcgaaa	cccgcagga	ctataaagat	540
accaggcgtt	tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	600
ccgataacct	gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	660
gtaggtatct	cagttcggtg	taggtcgttc	gctccaagct	gggctgtgtg	cacgaacccc	720
ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	aaccggtaa	780
gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggtatg	840
taggcggtgc	tacagagttc	ttgaagtggg	ggcctaacta	cggctacact	agaaggacag	900
tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	960
gatccggcaa	acaaaccacc	gctggtagcg	gtggtttttt	tgtttgcaag	cagcagatta	1020
cgcgcaaaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	1080
agtggaacga	aaactcacgt	taagggtatt	tggtcatgag	attatcaaaa	aggatcttca	1140
cctagatcct	tttaaattaa	aatgaagtt	ttaaatcaat	ctaaagtata	tatgagtaaa	1200
cttggtctga	cagttaccaa	tgcttaatca	gtgaggcacc	tatctcagcg	atctgtctat	1260
ttcgttcatc	catagttgcc	tgactccccg	tcgtgtagat	aactacgata	cgggagggct	1320
taccatctgg	ccccagtgct	gcaatgatac	cgcgagaccc	acgctcaccg	gctccagatt	1380
tatcagcaat	aaaccagcca	gccggaagg	ccgagcgcag	aagtggctct	gcaactttat	1440
ccgcctccat	ccagtctatt	aattgttgcc	gggaagctag	agtaagtagt	tcgccagtta	1500
atagtttgcg	caacgttggt	gccattgcta	caggcatcgt	ggtgtcacgc	tcgtcgtttg	1560
gtatggcttc	attcagctcc	ggttcccaac	gatcaaggcg	agttacatga	tcccccatgt	1620
tgtgcaaaaa	agcgggttagc	tccttcggtc	ctccgatcgt	tgtcagaagt	aagttggccg	1680
cagtgttatc	actcatgggt	atggcagcac	tgcataattc	tcttactgtc	atgccatccg	1740
taagatgctt	ttctgtgact	ggtgagtact	caaccaagtc	attctgagaa	tagtgtatgc	1800
ggcgaccgag	ttgctcttgc	ccggcgtaaa	tacgggataa	taccgcgcca	catagcagaa	1860
ctttaaagt	gtcatcatt	ggaaaacgtt	cttcggggcg	aaaactctca	aggatcttac	1920
cgctgttgag	atccagttcg	atgtaaccca	ctcgtgcacc	caactgatct	tcagcatctt	1980
ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	gcaaaatgcc	gcaaaaaagg	2040
gaataaggcg	gacacgaaaa	tggtgaatac	tcatactctt	cctttttcaa	tattattgaa	2100
gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	2160
aacaaatagg	ggttccgcgc	acatttcccc	gaaaagtgcc	acctgacgcg	ccctgtagcg	2220
gcgcattaag	cgcggcgggt	gtggtggtta	cgcgcagcgt	gaccgctaca	cttgccagcg	2280
ccctagcgcc	cgctcctttc	gctttcttcc	cttcctttct	cgccacgttc	gccggctttc	2340
cccgtcaagc	tctaaatcgg	gggctccctt	tagggttccg	atttagtgct	ttacggcacc	2400
tcgaccccaa	aaaacttgat	tagggtgatg	gttcacgtag	tgggccatcg	ccctgataga	2460
cggtttttcg	ccctttgacg	ttggagtcca	cgttctttta	tagtggactc	ttggttccaaa	2520
ctggaacaac	actcaaccct	atctcggctc	attcttttga	tttataaggg	attttgccga	2580
tttcggccta	ttggttaaaa	aatgagctga	tttaacaaaa	atttaacgcg	aattttaaca	2640
aaatattaac	gcttacaatt	tccattcgcc	attcaggctg	cgcaactgtt	gggaaggcg	2700
atcgggtcgg	gcctcttcgc	tattacgcca	gctggcgaaa	gggggatgtg	ctgcaaggcg	2760
attaagttgg	gtaacgccag	ggttttccca	gtcacgacgt	tgtaaaacga	cggccagtga	2820
attgtaatac	gactcactat	aggcgcaatt	ggagctcgtc	gagaccagat	gttttacact	2880

tgaccgtaaa tgagcaccctg aagaaaccgg tcacattcat ttcgaagggtg gagaaagcgg 2940  
aagatgactc aaacaagtaa tcggttgatg ttcgtcagtt catgtcactc ctatgaagga 3000  
gtcaagttca aaatggtatg ttgagtttca aactttttatg cttaaactttt tttctttatt 3060  
ttcgtaata atggaagaga accaattctc ttgtatctaa agattatcca tctatcatcc 3120  
aatttgagtg ttcaattctg gatgttggtg taccctacat tctacaacca tgtagccaat 3180  
tattatgaat ctggctttga tttcagttgt gttcttttct tttttttctt tgcataattg 3240  
catttagaat gttaataat taagttactg tatttccaca tacattagtt ccaagaatat 3300  
acatatatta atttattttt cttaaaaatg ttttggaatg actaatattg acaacgaaaa 3360  
tagaagctat gctaaaccat tacgtatatg tgacttcaca tgttggtgtt ttacattccc 3420  
tatatatatg gatggctgtc acaatcagaa acgtgatcga aaaaagacaa acagtgtttg 3480  
cataaaaaaga ctatttcgtt tcattgacaa tttgtgttta tttgtaaaga aaagtggcaa 3540  
agtggaaatt gagttcctgc aagtaagaaa gatgaaataa aagacttgag tgtgtgtttt 3600  
tttcttttat ctgaaagctg caatgaaata ttcctaccaa gcccgtttga ttattaattg 3660  
gggtttggtt ttcttgatgc gaactaattg gttatataag aaactatata atccatgtta 3720  
attcaaaaat tttgatttct cttgtaggaa tatgatttac tatatgagac tttcttttctg 3780  
ccaataatag taaatccaaa gatatttgac cggaccaaaa cacattgatc tatttttttag 3840  
tttattttaat ccagtttctc tgagataatt cattaaggaa aacttagtat taacccatcc 3900  
taagattaaa taggagccaa actcacattt caaatattaa ataacataaa atggatttaa 3960  
aaaaatctata cgtcaaat ttttatgac atttcttatt taaatttata tttaatgaaa 4020  
tacagctaag acaaaccaaaa aaaaaaatac tttctaagtg gtccaaaaca tcaattccgt 4080  
tcaatattat taggtagaat cgtacgacca aaaaaaggta ggtaataacg aattagaaac 4140  
atatctataa catagtatat attattacct attatgagga atcaaaatgc atcaaatatg 4200  
gatttaagga atccataaaa gaataaaatc tacgggaaaa aaaatggaat aaattctttt 4260  
aagtttttta tttgtttttt atttggtagt tctccatttt gttttatttc gtttggtttt 4320  
atttgtgtcca aatactttgt aaaccaccgt tgtaattctt aaacgggggtt ttcacttctt 4380  
ttttatattc agacataaag catcggtggt tttaatcaat caatagattt tatttttctt 4440  
ctcaattatt agtaggtttg atgtgaactt tacaaaaaaa acaaaaacaa atcaatgcag 4500  
agaaaaagaaa ccacgtgggc tagtcccacc ttgtttcatt tccaccacag gttcgatctt 4560  
cgttaccgtc tccaatagga aaataaacgt gaccacaaaa aaaaaacaaa aaaaagtcta 4620  
tataattgctt ctctcaagtc tctgagtgct atgaaccaa gtaaaaaaca aagactcgac 4680  
ctgcaggcat gcaagcttat cgtcgactac gtaagtttct gcttctacct ttgatata 4740  
tataataatt atcattaatt agtagtaata taatatttca aatatttttt tcaaaaataa 4800  
agaatgtagt atatagcaat tgcttttctg tagtttataa gtgtgtatat ttttaatttat 4860  
aacttttcta atatatgacc aaaatttgtt gatgtgcagg tatcaccgga tccatcgaat 4920  
tcggtacgct gaaatcacca gtctctctct acaaatctat ctctctctat tttctccata 4980  
aataatgtgt gagtagtttc ccgataaggg gaatttaggg ttcttatagg gtttcgctca 5040  
tgtgttgagc atataagaaa cccttagtat gtatttgat ttgtaaaata ctctatcaa 5100  
taaaatttct aattcctaaa accaaaatcc agtactaaaa tccagatctc cttaaagtccc 5160  
tatagatctt tgctgtgaat ataaaccaga cacgagacga ctaaacctgg agcccagacg 5220  
ccgttcgaag ctagaagtac cgcttaggca ggaggccgtt agggaaaaga tgctaaggca 5280  
gggttggtta cggtgactcc cccgtagggt ttggtttaa atgatgaagt ggacggaagg 5340  
aaggaggaag acaaggaaag ataagggtgc aggcctgtg caaggtaaga agatggaaat 5400  
ttgatagagg tacgctacta tacttatact atacgctaag ggaatgcttg tatttatacc 5460  
ctataccccc taataacccc ttatcaattt aagaaataat ccgcataagc ccccgcttaa 5520  
aaattggtat cagagccatg aataggctca tgaccaaacc tcaagaggat aaaacctcac 5580  
caaaatacga aagagttctt aactctaaag ataaaagatc tttcaagatc aaaactagtt 5640  
ccctcacacc ggtgacgggg atcgcgatgg gtac 5674

pls explain  
"N" location.

See error  
explanation on  
page 7.

The type of errors shown exist throughout  
the sequence listing. Please check subsequent  
sequences for similar errors.

## VARIABLE LOCATION SUMMARY

DATE: 02/17/2005

PATENT APPLICATION: US/10/522,341

TIME: 12:04:45

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\02172005\J522341.raw

Use of n's or Xaa's (NEW RULES):

Use of n's and/or Xaa's have been detected in the Sequence Listing.

Use of <220> to <223> is MANDATORY if n's or Xaa's are present.in <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

Error Explanation: 2

Seq#:55; N Pos. 5014

Seq#:57; N Pos. 6697

Seq#:59; N Pos. 1026

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/522,341

DATE: 02/17/2005

TIME: 12:04:45

Input Set : A:\sequence listing.txt

Output Set : N:\CRF4\02172005\J522341.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:228 M:283 W: Missing Blank Line separator, <220> field identifier  
L:231 M:283 W: Missing Blank Line separator, <220> field identifier  
L:236 M:283 W: Missing Blank Line separator, <220> field identifier  
L:354 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:4 ✓  
L:357 M:200 E: Mandatory Header Field missing, <220> Tag not found for SEQ ID#:4 ✓  
L:443 M:283 W: Missing Blank Line separator, <220> field identifier  
L:641 M:283 W: Missing Blank Line separator, <220> field identifier  
L:645 M:283 W: Missing Blank Line separator, <400> field identifier  
L:866 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1092 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1199 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1384 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1568 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1713 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1802 M:283 W: Missing Blank Line separator, <220> field identifier  
L:1891 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2014 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2260 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2606 M:283 W: Missing Blank Line separator, <220> field identifier  
L:2834 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3185 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3413 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3642 M:283 W: Missing Blank Line separator, <220> field identifier  
L:3754 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:37  
L:3847 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4039 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4225 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4411 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4597 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4782 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4802 M:283 W: Missing Blank Line separator, <220> field identifier  
L:4807 M:252 E: No. of Seq. differs, <211> LENGTH:Input:27 Found:21 SEQ:50 ✓  
L:4814 M:283 W: Missing Blank Line separator, <220> field identifier ✓  
L:4819 M:252 E: No. of Seq. differs, <211> LENGTH:Input:26 Found:20 SEQ:51 ✓  
L:4826 M:283 W: Missing Blank Line separator, <220> field identifier ✓  
L:4846 M:283 W: Missing Blank Line separator, <220> field identifier ✓  
L:4851 M:252 E: No. of Seq. differs, <211> LENGTH:Input:27 Found:20 SEQ:53 ✓  
L:4857 M:283 W: Missing Blank Line separator, <220> field identifier ✓  
L:4869 M:283 W: Missing Blank Line separator, <220> field identifier ✓  
L:4957 M:258 W: Mandatory Feature missing, <221> Tag not found for SEQ ID#:55 ✓  
L:4957 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:55 ✓  
L:4957 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:4980 ✓  
L:4975 M:283 W: Missing Blank Line separator, <220> field identifier ✓  
L:5087 M:283 W: Missing Blank Line separator, <220> field identifier ✓  
L:5203 M:258 W: Mandatory Feature missing, <221> Tag not found for SEQ ID#:57 ✓  
L:5203 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:57 ✓

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/522,341

DATE: 02/17/2005

TIME: 12:04:45

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\02172005\J522341.raw

L:5203 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57 after pos.:6660  
L:5262 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5510 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5597 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:59 after pos.:1011  
L:5597 M:254 E: No. of Bases conflict, LENGTH:Input:325 Counted:1032 SEQ:59  
L:5597 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:6  
L:5597 M:112 C: (48) String data converted to lower case,  
L:5597 M:252 E: No. of Seq. differs, <211> LENGTH:Input:1011 Found:1032 SEQ:59  
L:5673 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5761 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5842 M:283 W: Missing Blank Line separator, <220> field identifier  
L:5923 M:283 W: Missing Blank Line separator, <220> field identifier  
L:6003 M:283 W: Missing Blank Line separator, <220> field identifier  
L:6015 M:283 W: Missing Blank Line separator, <220> field identifier  
L:6027 M:283 W: Missing Blank Line separator, <220> field identifier